

Subsurface Characterization Using a Cellular Automaton Approach

H.M. Schuttelaars · F.M. Dekking · C. Berentsen

Received: 21 November 2007 / Accepted: 6 January 2009 / Published online: 18 June 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract In this paper, a random Cellular Automaton model is developed to characterise heterogeneity of geological formations. The CA-model is multilateral and can be easily applied in both two and three dimensions. We demonstrate that conditioning on well data is possible and the method is numerically efficient. To construct the model, the subsurface is subdivided into N cells, with an initial lithology assigned to each cell. Rules to update the current cell states are chosen from a set of rules, independently for each cell. The model converges typically in less than 50 iterations to a steady state or periodic solution. Within one period the realisations exhibit similar statistical properties. The final fraction of the various lithologies can be tuned by choosing the proper initial fractions. In this way, geological knowledge of those fractions can be satisfied.

Keywords Transition probability · Two and three dimensional characterisation · Well-data

H.M. Schuttelaars (✉) · F.M. Dekking
Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4,
2628 CD Delft, The Netherlands
e-mail: h.m.schuttelaars@tudelft.nl

C. Berentsen
Department of Earth Sciences, University of Utrecht, P.O. Box 80021, 3508 TA Utrecht,
The Netherlands

C. Berentsen
Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1,
2628 CN Delft, The Netherlands

C. Berentsen
Shell International Exploration and Production, Kessler Park 1, 2288 GS Rijswijk, The Netherlands

1 Introduction

It is well-known that geological formations exhibit significant spatial heterogeneity. This variability manifests itself at various scales. A precise description of the configuration (geometry), behaviour, and properties of various geological patterns in the underground is essential in many engineering fields, such as hydrogeology, civil and geo-technical engineering, safety assessment studies, study of enhanced oil recovery processes, etc. (for a recent discussion on the importance of subsurface characterisation, see de Marsily et al. 2005). Numerous methods to characterise the subsurface are reported in the literature that utilise hard data (borehole data) and/or soft data. These methods can be broadly classified as either process-based or stochastic approaches (for a comprehensive overview, see de Marsily et al. 2005). In the former type of approach, one tries to model complex processes resulting in the formation of observed depositional sequences. However, this approach is very difficult to condition on existing hard data. Thus, the model results are not necessarily in agreement with existing borehole data. Furthermore, this method requires data that is often not available, e.g. sedimentation and erosion rates, turbidity factors, etc. The approach is also inefficient in terms of computational time and storage, it is highly sensitive to initial and boundary conditions, and interpretation and validation of results are subjective. Stochastic models are based on geostatistical theory. Examples of these models are stochastic Gaussian fields, Indicator-based models (Sequential Indicator Simulation), object-based models (Boolean models), transition probability-based models (e.g. Markov chains), and simulation methods using multi-point statistics (Strebelle 2002). The stochastic approach is in general more appealing when compared to the process-based approach in the sense that it honours known data and their locations and patterns of spatial variability.

Elfeki and Dekking (2001, 2005) developed a stochastic methodology to characterise subsurface heterogeneity based on a two-dimensional extension of the one-dimensional Markov chain theory. Their method is efficient and can be conditioned on any number of boreholes. One of the main shortcomings of this model is its unilaterality, i.e. information is propagated in only one direction. This also leads to a conceptual difficulty when extending it to three dimensions. Therefore, the main aim of this paper is to develop a new model formulation that can be easily extended to three spatial dimensions. A natural extension of the Markov chain method that satisfies these requirements is the Cellular Automata (CA) method. The characteristic behaviour of the new model will be tested in two dimensions, using an idealised geometry. It will be shown that the CA model can be easily conditioned on available data and is computationally efficient. The structure of the paper is as follows: First, the CA model is introduced and the implementation of the CA algorithm is described, then demonstration of the model's effectiveness is presented. In the final section, we conclude with a discussion of the model and its results.

2 Inhomogeneous and Random Cellular Automata

CA have been studied since the 1950s (Ulam 1952), and have recently been placed on a pedestal by Wolfram (2002) in his book “A new kind of science”. However, in this

paper we shall be concerned with random CA, which only appear marginally in Wolfram's book. Random CA should not be confused with probabilistic CA (see Toom 1995), where the rules are updated randomly at every iteration step. In random CA, the rules are chosen once, randomly, and the system runs with these rules for all iterations—starting from a random initial configuration. Two state random CA was introduced by Kauffman (1969, 1984) as a model for complex genetic regulatory systems. In his model, there is extra randomness: The cells are nodes in a random directed graph, where each cell has a mother and a father randomly chosen from the other cells. In our application the graph is part of the D -dimensional integer lattice, and the parents of a cell are simply its neighbours. The basic idea of the working of a CA is that every cell has a rule φ which inputs the states of the neighbouring cells and outputs the new state of the cell. In the following, we will formalise these concepts. Before we discuss random CAs, we have to introduce inhomogeneous CAs, since the realisations of random CAs are inhomogeneous CAs. Inhomogeneity refers to the possibility that different cells have different rules. Let N be the total number of cells in the CA. In the following, we will, for simplicity of notation, consider the 1-dimensional case, where the cells are the integers $1, 2, \dots, N$, and where cell i has neighbours $i - 1$ and $i + 1$. The state of cell i is given by $x(i)$, where $x(i)$ is an element of a finite set of states which in this case represent the lithology. The rule to update cell i is denoted by φ_i .

The states of the automaton can be updated in various ways. The most straightforward way is by updating all the cells simultaneously:

$$x_i(n+1) = \varphi_i(x_{i-1}(n), x_{i+1}(n)) \quad \text{for } i = 2, \dots, N-1.$$

However, this method often results in so-called checkerboard patterns. One way to avoid these patterns is to first update only those cells for which the cell number is even, leaving the cells characterised by an odd cell number unchanged (Fig. 1). This results in an intermediate state that is used to update the odd cells, leaving the even cells unchanged.

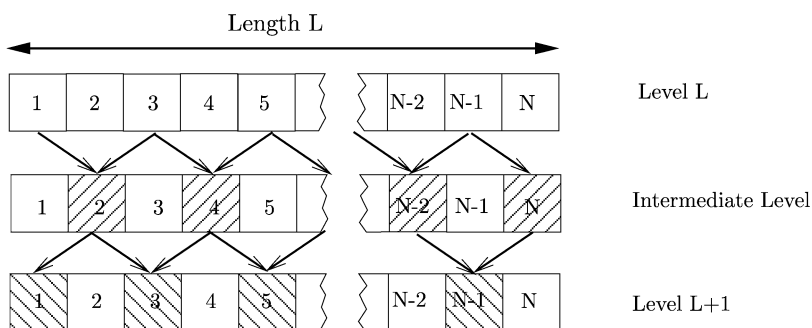


Fig. 1 A one-dimensional layer of length L is subdivided into N cells. Here N is chosen to be even. Based upon the lithologies of its surrounding cells, the type of lithology of every cell is updated. First, only the states of the even cells are updated, resulting in an intermediate level. Subsequently, the odd cells of the intermediate level are updated, resulting in a complete update of all cells

This procedure will be used in the remainder of this article. To deal with the end points, we shall update the states in cell 1 and cell N by a rule that only depends on one variable:

$$x_1(n+1) = \varphi_1(x_2(n)), \quad x_N(n+1) = \varphi_N(x_{N-1}(n)).$$

In a random CA, the rules φ_i are chosen according to a distribution Φ on the set of rules, independently for each cell i . Also, the initial state is a sequence of independent random variables $X_i(0)$, each having the same distribution on the set of states, independent from the (Φ_i) . The state of cell i at time n is now a random variable, denoted $X_i(n)$. Using the simplest update strategy, the value of the i th cell at iteration $n+1$ is given by

$$X_i(n+1) = \Phi_i(X_{i-1}(n), X_{i+1}(n)),$$

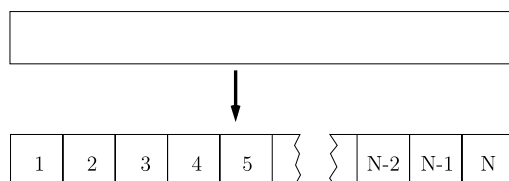
but we will use the checkerboard update scheme. In the next section, we introduce the working of a random CA in the context of subsurface characterisation.

3 The CA Model

To model the subsurface lithologies using CA, the first step is to divide the subsurface into a number of cells N (Fig. 2). Every cell is identified by a vector $\mathbf{i} = (i_1, \dots, i_d, \dots, i_D)$ where i_d is the number of the cell in the d th direction. Here D is the spatial dimension under consideration.

As a next step, one has to decide how to update a given cell \mathbf{i} , i.e. one has to choose a CA-rule for cell \mathbf{i} . First, choose the stencil of cells whose information is used to update cell \mathbf{i} . In principle, the neighbourhood of cell \mathbf{i} may consist of any number of cells, here we restrict ourselves to the direct neighbourhood of cell \mathbf{i} . Hence in one dimension, the neighbourhood only contains the cell right or east (denoted by e) and left or west (w) of it, in two dimensions the neighbourhood also contains the cells above or north of cell \mathbf{i} , denoted by n , and south (s) of cell \mathbf{i} . In three dimensions, the cell in front of (f) and behind (b) cell \mathbf{i} are considered as well. From now on $\mathcal{N}(\mathbf{i}) = \{\mathbf{i}_e, \mathbf{i}_w\}$, $\mathcal{N}(\mathbf{i}) = \{\mathbf{i}_n, \mathbf{i}_e, \mathbf{i}_s, \mathbf{i}_w\}$ and $\mathcal{N}(\mathbf{i}) = \{\mathbf{i}_n, \mathbf{i}_e, \mathbf{i}_s, \mathbf{i}_w, \mathbf{i}_f, \mathbf{i}_b\}$ in one, two and three dimensions, respectively (Fig. 3). Furthermore, $\mathbf{i}_c \in \mathcal{N}(\mathbf{i})$ denotes the location of any cell neighbouring \mathbf{i} . The states of the cells are the lithologies of the subsurface. The model distinguishes r lithologies from a set $\mathcal{S} = \{1, 2, \dots, r\}$. Given the lithologies of the neighbouring cells, the lithology in the cell under consideration will be updated according to a CA-rule φ .

Fig. 2 The subdivision of a one-dimensional stretch of subsurface by N cells



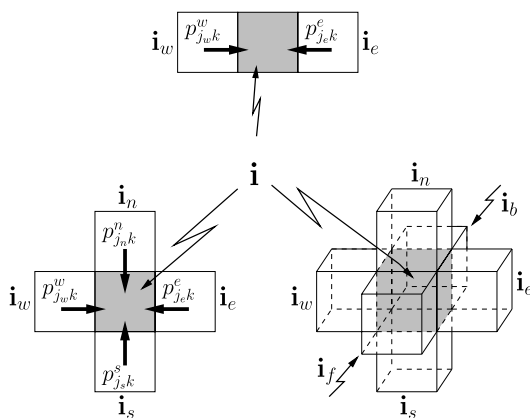


Fig. 3 The subdivision of the subsurface in dimension one (*top figure*), two (*figure bottom left*) and three (*figure bottom right*). In these figures, the naming convention for the cells adjacent to the cell under consideration (cell \mathbf{i}) is given. In the one and two dimensional cases, the transition probabilities from cell \mathbf{i}_c with state j_c to cell \mathbf{i} with state k are given by the transition probabilities $p_{j_c k}$ and indicated by the arrows. Here $c \in \{e, w\}$ in one dimension, $c \in \{n, e, s, w\}$ in two and $c \in \{n, e, s, w, f, b\}$ in three dimensions

To obtain CA-rules that may reproduce the subsurface structure of a given site, we need the probabilities

$$p_{\vec{j},k} \quad (1)$$

that a given cell \mathbf{i} has a specific state $k \in \mathcal{S}$, given the state of the neighbouring cells $\vec{j} \in \mathcal{S}^{2D}$. Hence $\vec{j} = (j_e, j_w)$ in one dimension, $\vec{j} = (j_n, j_e, j_s, j_w)$ in two dimensions, and $\vec{j} = (j_n, j_e, j_s, j_w, j_f, j_b)$ in three dimensions, with $j_c \in \mathcal{S}$.

To obtain an estimate of $p_{\vec{j},k}$ (denoted by $\hat{p}_{\vec{j},k}^c$) one-sided transition probabilities are used. The one-sided probabilities give the probability that, given that the neighbouring cell $\mathbf{i}_c \in \mathcal{N}(\mathbf{i})$ has a specific lithology j_c , cell \mathbf{i} has lithology k . Following Elfeki and Dekking (2001), using the one-sided probabilities, the required probabilities (1) can be constructed in any dimension:

$$\hat{p}_{\vec{j},k}^c = \frac{\prod_{c \in \mathcal{N}(\mathbf{i})} \hat{p}_{j_c,k}^c}{\sum_{k'} \prod_{c \in \mathcal{N}(\mathbf{i})} \hat{p}_{j_c,k'}^c}. \quad (2)$$

The use of one-sided probabilities corresponds to the simplifying assumption that the various one-sided transition probabilities are independent. One-sided probabilities are given by

$$\hat{p}_{j_c,k}^c = \frac{n_{j_c,k}}{\sum_{k'} n_{j_c,k'}}, \quad (3)$$

where $n_{j_c,k}$ is the number of times that lithology k appears in cell \mathbf{i} with lithology j_c in cell \mathbf{i}_c . The numbers $n_{j_c,k}$ can be obtained from borehole data, geological information and training images. For example, considering two states, these one-sided transition

probabilities can be written down as

$$\hat{P}^c = \begin{bmatrix} \lambda_1^c & 1 - \lambda_1^c \\ 1 - \lambda_2^c & \lambda_2^c \end{bmatrix}, \quad (4)$$

where $\lambda_k^c = \hat{P}_{kk}^c$ denotes the probability that given that cell \mathbf{i}_c has state k , with $k \in \{1, 2\}$, cell \mathbf{i} has state k as well. From now on, λ_k^c will be called the persistence of lithology k in the c -direction: If the persistence is 1, both cells have the same lithology; if the persistence is 0, they will certainly have a different lithology.

Using the probabilities $\hat{p}_{\vec{j},l}$, a standard simulation procedure is employed to update the cell states: for every neighbourhood state \vec{j} of cell \mathbf{i} a number $u^{\vec{j}}$ is chosen randomly from $[0, 1]$. For each neighbourhood state \vec{j} , the state of cell \mathbf{i} is updated to lithology $k^* \in \mathcal{S}$ for which

$$\sum_{l=1}^{l=k^*-1} \hat{p}_{\vec{j},l} < u^{\vec{j}} \leq \sum_{l=1}^{l=k^*} \hat{p}_{\vec{j},l}.$$

This procedure is repeated for every possible neighbourhood state \vec{j} . In this way, the CA-rule, which is fixed during the whole iteration procedure, is defined.

In Fig. 4, a small, two-dimensional example is shown. The subsurface is divided into 5×5 cells and all the one-sided transition probabilities are equal to 0.55. The lithology that a cell will obtain after an iteration depends on the states of the cells in its neighbourhood. As an example, the neighbourhood of the central cell is indicated by a thick contour. In this experiment, the CA reaches its attractor after seven iterations. The global state obtained after the seventh iteration step is the first global state of the periodic attractor. The periodic attractor has a period four: the global state after $n + 4$ iterations is identical to the one after n iterations (with $n > 4$).

4 Algorithm

The procedure to find one realisation using the method discussed in the previous section is as follows:

1. The one, two, or three-dimensional domain is discretised. In principle, this discretisation can result in cells of any shape. In this paper, however, we will only use rectangular cells and focus on two-dimensional domains.
2. Well data is inserted in the cells that represent the locations of the wells in the CA-model. The type of lithology in these cells is not changed during the update procedure as described in Steps 5–6.
3. Every cell is assigned an initial lithology type.
4. A CA-rule is assigned to every cell. This rule defines how the lithology of a specific cell will be updated, given the lithologies in a neighbourhood of this cell. The CA-rules are kept fixed during the update procedure in Steps 5–6.

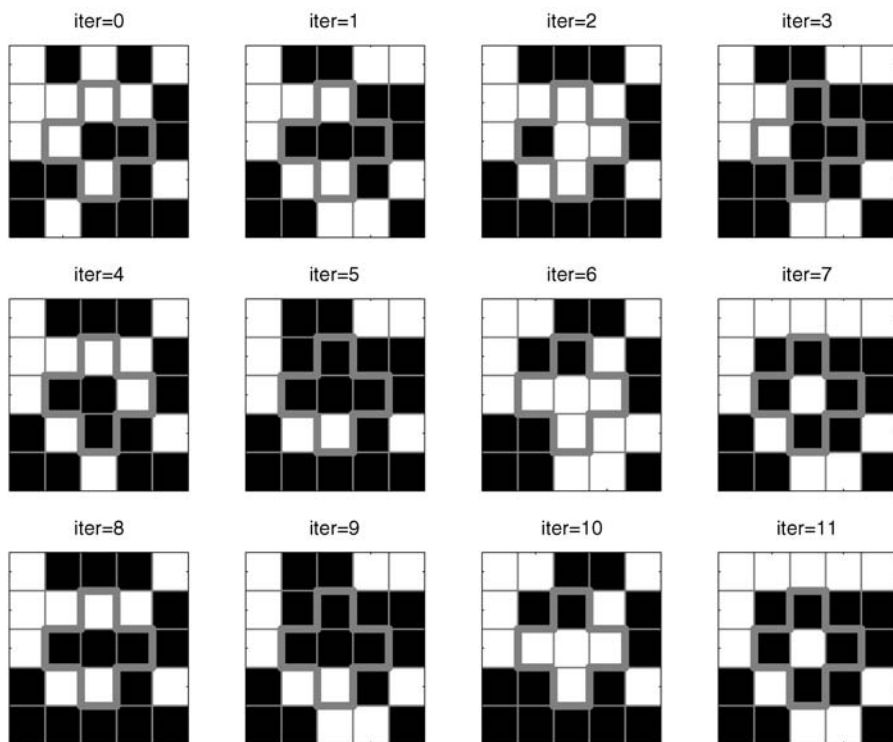


Fig. 4 A small, two-dimensional CA with period four that reaches its periodic attractor after seven iterations. For a more detailed description, see text

5. Using the given distribution of lithologies and the CA-rules, defined in Step 3, the lithologies are updated according to the so-called checkerboard scheme: First, only those cells $\mathbf{i} = (i_1, \dots, i_D)$, with D the spatial dimension under consideration, are updated for which the sum of the cell coordinates $\sum_{l=1}^D i_l$ is an even number and the cells characterised by an odd cell number are unchanged. This results in an intermediate state that is used to update the odd cells, leaving the even cells unchanged.
6. The previous step, Step 5, is repeated until either the configuration of the subsurface does not change anymore (a steady global state is reached), or a periodic solution with period q larger than one is found, updating the lithological distribution in the subsurface will result in the same subsurface structure as observed q update steps before.
7. If a steady state is reached, this solution is given as one realisation characterising the subsurface. In case of a periodic solution, the first global state is chosen as realisation. Its representativeness is investigated by comparing this state to other states in the periodic solution.

5 Model Results

5.1 Two Dimensional CA

In this section, two-dimensional examples will be discussed to illustrate the working and sensitivities of the proposed CA-method. In the first series of experiments, we will only consider two active sediment types, and the chosen transition probabilities are synthetic. The lithologies in the first and last column are synthetic and do not change in the experiments, since they correspond to the two wells. Next, the vertical probabilities are determined from information obtained from 2 wells. In all experiments, use is made of probabilities that are defined by (2) to obtain the CA-rules.

5.2 Synthetic Two-dimensional Experiments

To study the behaviour of the CA-model, five synthetic experiments are performed (see Table 1 for a short overview of the experiments). In all these experiments, the subsurface is divided into $N_X = 100$ by $N_Y = 100$ cells. Using a standard simulation technique, a fraction Γ of all cells is assigned lithology 0. The parameter Γ is varied between 0.3 and 0.7 with steps of 0.05. The one-sided transition probabilities \hat{P}^c with $c \in \{n, e, s, w\}$ are prescribed and chosen to be identical for both sediment types. In every experiment, the one-sided transition probabilities and the fraction Γ are varied. For every parameter setting, 250 realisations are made to get information about the mean and standard deviation of the cycle length, the number of iterations necessary to reach the final solution, the fraction of the two lithologies and the fraction of cells that changes within the periodic attractor. For every realisation, one can choose to either start with a new initial condition, or determine new CA-rules, or do both.

Experiment E1 The one-sided transition probabilities are identical in all directions and are varied from 0.75 to 0.999 with steps of 0.05. The model domain consists of a rectangle with two active lithologies. The first and last column of the model domain represent the locations of the wells and the prescribed lithologies in these columns do no change. In Fig. 5, four different realisations are shown. They were obtained with the parameter settings of the reference experiment (E1 in Table 1). These four realisations clearly show the stochastic variability. The persistences were equal in all directions and set to 0.8 for both lithologies. Initially Γ was 0.5. The

Table 1 Parameter settings for the various experiments

Exp	$\hat{P}^e, \hat{P}^s, \hat{P}^w, \hat{P}^n$	Default Γ	CA-rules fixed	Initial Condition fixed	Domain
E1	Identical	0.5	No	No	Rectangular
E2	Identical	0.5	Yes	No	Rectangular
E3	Identical	0.5	No	Yes	Rectangular
E4	Identical	0.5	No	No	General
E5	Not identical	0.5	No	No	Rectangular

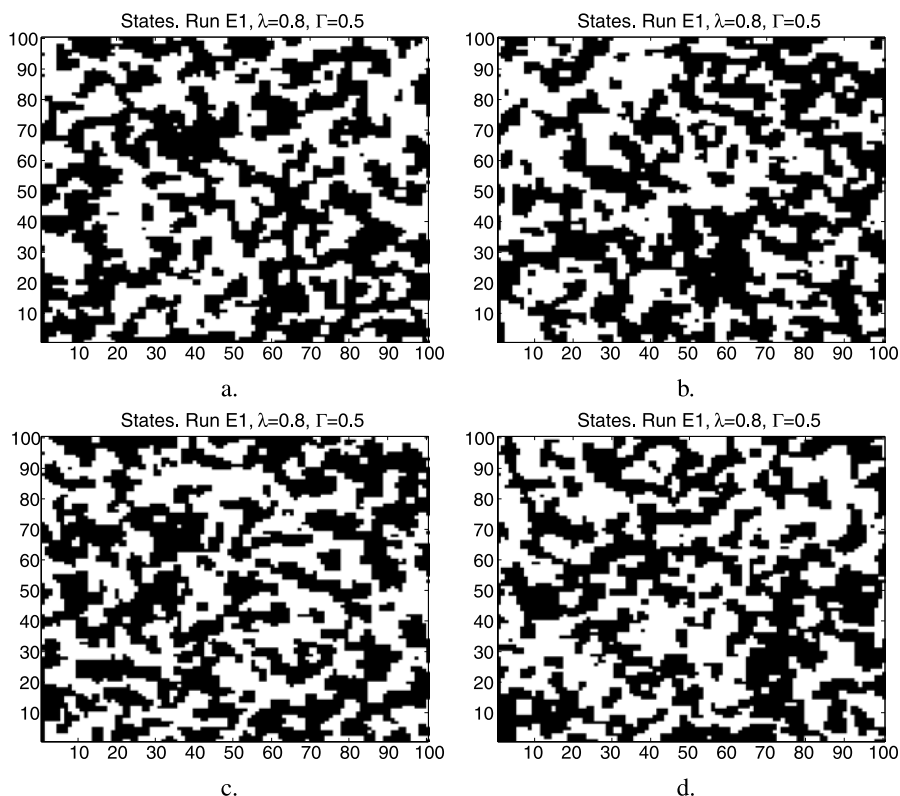


Fig. 5 Four different global realisations obtained with the parameter settings of experiment **E1**

converged subsurface characterisations obtained with the CA-method will in general be periodic. In this case, we show the first global state in this periodic attractor.

However, Fig. 5 clearly shows that global states related to different periodic solutions (i.e. obtained with different initial conditions and/or different CA-rules) may vary significantly. Therefore, statistical properties of the global states are given in Fig. 6.

In Fig. 6(a), the average cycle length of the periodic attractor (i.e. the number of update steps necessary to visit all states of the periodic attractor) and its standard deviation are shown for various persistences and initial fractions Γ . If the persistence is close to one, most of the attractors are steady state solutions and only a small fraction of the global states has a short period. Decreasing the persistences results in an increase of the cycle length. In Fig. 6(b), the average number of iterations needed to reach the periodic attractor and its standard deviation are shown. The higher the persistence, the faster the periodic attractor is reached. When the persistence is larger than ~ 0.9 , approximately 10 to 20 iterations are needed to reach the attractor, for transition probabilities in the order of 0.75, on average approximately 50 iterations are necessary.

In Fig. 6(c), the fraction of lithology 0 after convergence, Γ_{final} , is plotted as a function of the initial occurrence of lithology 0, Γ_{init} . A Γ_{init} of 0.5 results in a

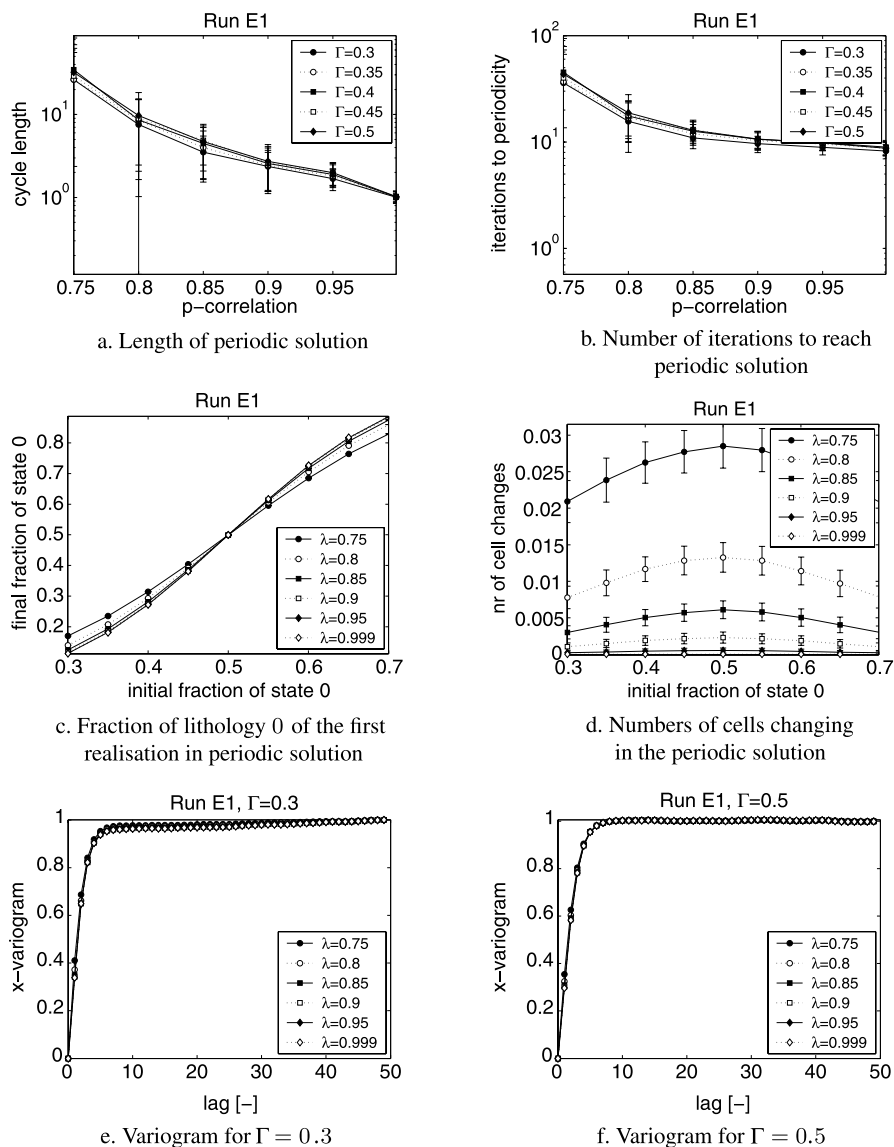


Fig. 6 Statistical properties of the experiments E1

mean Γ_{final} of 0.5 as well. The standard deviation is approximately 10%. If the initial fraction is initially larger (smaller) than 0.5, the resulting fraction is larger (smaller) than the initial fraction. In Fig. 6(d), the fraction of the cells that change lithology within the periodic attractor is plotted. It is clear that the overall patterns within a periodic attractor do not change much. Hence, any realisation taken from the periodic attractor is representative. In Fig. 6(e) and (f) the normalised semi-variogram $\gamma(L)$ in the horizontal direction is shown for $\Gamma_{\text{init}} = 0.3$ and $\Gamma_{\text{init}} = 0.5$, respectively, for

different values of the persistence. This quantity is calculated as

$$\gamma(L) = \frac{1}{2\sigma_s^2} \left[\frac{1}{N_Y(N_X - L)} \sum_{j=1}^{N_Y} \sum_{i=1}^{N_X-L} (S_{i,j} - S_{i+L,j})^2 \right],$$

with σ_s^2 the state variance and $S_{i,j}$ is the lithology of cell (i, j) . If $\gamma(L) = 0$, the state is fully correlated over a distance L , when $\gamma(L) = 1$ the state is fully uncorrelated. Figure 6(f) shows that for $L = 10$ the system is uncorrelated when $\Gamma_{\text{init}} = 0.5$. If Γ_{init} decreases or increases, the lag for which the system becomes fully uncorrelated slightly increases.

Experiments E2–E5 To illustrate the sensitivity of the model results to the choice of the CA-rules, the initial condition, the model domain and the one-sided transition probabilities, four more experiments were performed. To investigate the model sensitivity with respect to the choice of the CA-rules, the CA-rules are not varied from realisation to realisation in experiment **E2**. This is the only difference between experiment **E2** and **E1**. Experiment **E3** only differs from experiment **E1** in the choice of the initial condition. In experiment **E3**, the initial condition is the same for all realisations. In experiment **E1**, the initial condition is different for every realisation. In experiment **E4**, the domain is varied: Instead of a rectangular domain, a more general domain is considered bounded from the top by air and from the bottom by bedrock. The position of the top and bottom boundaries is prescribed and does not change. No information is propagated from these boundaries into the domain. In experiment **E5**, the probabilities in the vertical direction will be chosen different from those in the horizontal, i.e. $\hat{p}^e = \hat{p}^w$ and $\hat{p}^n = \hat{p}^s$, but $\hat{p}^e \neq \hat{p}^n$.

Examples of the results obtained with the settings of experiments **E2–E5** are shown in Fig. 7. The patterns obtained with experiments **E2–E4** are quite similar. Note that in experiment **E4** four lithologies were considered: Apart from the two lithologies taken in account in the other experiments, the bedrock (dark gray) and the air (light gray) were taken into account as well. The patterns resulting from experiment **E5** are extended more in the horizontal direction compared to those of the other experiments. In Fig. 7(d), the horizontal persistence is larger than the vertical persistence (0.8 and 0.7, respectively) with $i \in \{1, 2\}$.

Again, realisations of the subsurface show a lot of variation. Therefore, the influence of the model variations on the statistical properties of several descriptors is depicted in Fig. 8 and is discussed below.

- Cycle length of the periodic attractor (CL): Fig. 8(a) shows that a decrease of either λ_H , λ_V or both of the persistences results in an increase of the CL. In this figure, the average CL of the periodic solution is shown as a function of λ_H (x-axis) and λ_V (y-axis). In these experiments, $\Gamma = 0.4$. For experiments **E1–E4**, the horizontal persistences (denoted by λ_H) and the vertical ones (λ_V) are identical; these results are depicted by four, almost identical lines in Fig. 8(a). When considering experiment **E5**, λ_H and λ_V may vary independently, resulting in a surface of experimentally observed cycle-lengths.

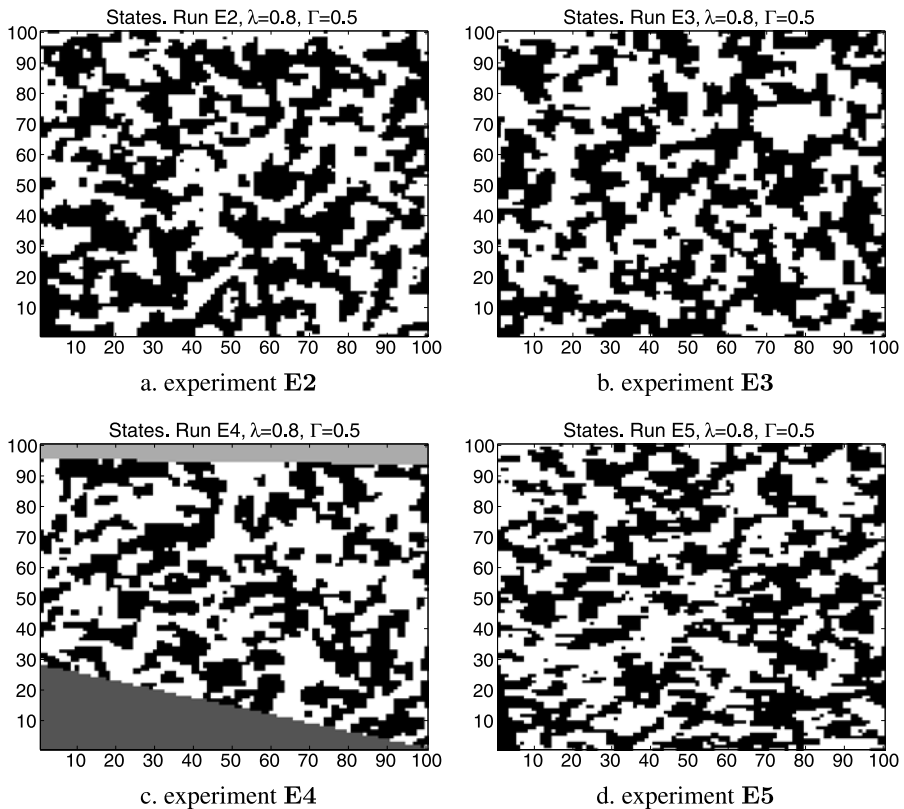


Fig. 7 One realisation for each of the four different experiments

- Number of iterations necessary to reach the periodic attractor (IP): In Fig. 8(b), a similar plot is shown for the average number of iterations needed to reach the periodic solution. The same trend is observed: Decreasing the persistences results in an increase of the number of iterations needed to reach the periodic solution.
- Final fraction of lithology 0 (FF): The observed fraction of lithology 0 in the first realisation of the periodic solution is smaller than its initial fraction when $\Gamma < 0.5$. Figure 8(c) illustrates this for $\Gamma = 0.4$. The higher the persistences, the smaller the resulting fraction. In the case of experiment E2, the resulting fractions deviate from those observed in the other experiments. The final fractions can be both larger and smaller than those found in the other experiments, depending on the (randomly chosen) CA-rules. Since the CA-rules are fixed in this experiment, it is even possible that Γ_{final} is larger than Γ_{init} although this is highly unlikely.
- Number of cells changing within the periodic attractor (CHG): Figure 8(d) shows that the number of cells that change within a periodic solution is small, from no cell changing for large persistences up to 6% for relatively small transition probabilities.
- Variogram: The variogram, shown in Fig. 8(e) rapidly increases from 0 to 1 when the lag is increased. It is almost independent of the vertical transition probabil-

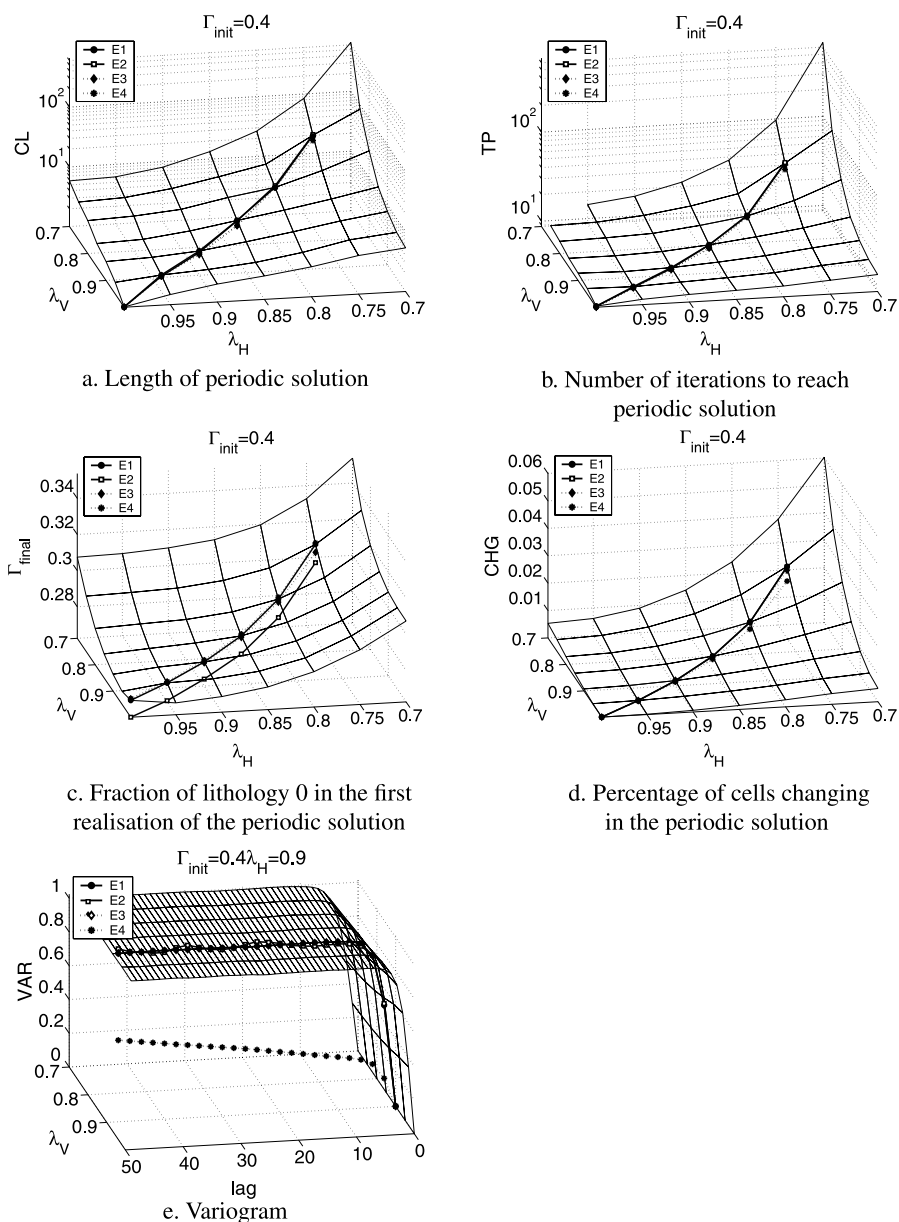


Fig. 8 Comparison of the statistical properties of the experiments **E1–E5**. The four lines in the different subfigures correspond to experiments **E1–E4**, the surface shows the results obtained with experiment **E5**

ity. The larger correlations observed in experiment **E4** are due to the cells with prescribed bedrock and air lithologies that cannot change. The variograms of the lithologies that are allowed to vary have the same structure as those obtained for the other experiments (not shown).

5.3 Using Well Information to Obtain the CA-rules

To demonstrate the use of well-data to obtain the transition probabilities, the Delaware River and its underlying aquifer system is taken as an example. The data used is collected from the literature (Navoy 1991) and extensively discussed in Elfeki and Dekking (2005).

In our example, we take wells 6 and 11 from the latter paper. In the first series of experiments, the various lithologies are grouped together in seven types and three major groups: sand (codes 1–3), clay (codes 4–6), and bedrock (Table 2). The last lithology can be either active or inactive. In the second series of experiments, all lithologies as described in Elfeki and Dekking (2005) are taken into account.

Two lithologies When grouping together the sandy lithologies and the clayey lithologies, it is found that in the well-logs considered there is a slight bias toward sand (colour-coded white in the figures). Hence, state transition bias is introduced in the persistences when the well-logs are subdivided into 100 equidistant cells. Using these cells, the following one-sided transition probabilities are obtained:

$$\hat{P}^n = \begin{bmatrix} 0.84 & 0.16 \\ 0.21 & 0.79 \end{bmatrix}, \quad \hat{P}^s = \begin{bmatrix} 0.82 & 0.18 \\ 0.20 & 0.8 \end{bmatrix} \quad (5)$$

and

$$\hat{P}^w = \hat{P}^e \begin{bmatrix} \lambda & 1 - \lambda \\ 1 - \lambda & \lambda \end{bmatrix} \quad (6)$$

with $\lambda \in \{0.75, 0.9999\}$. The top and bottom boundaries exhibit some topographic features but are fixed in this experiment. In Fig. 9(b), a realisation is depicted. This realisation shows that the fraction of lithology 0 (white) is larger than the fraction of lithology 1 (black). Figure 9(d) shows the final fraction of lithology 0 as a function of the initial fraction of lithology 0. To get $\Gamma_{\text{final}} > 0.5$, Γ_{init} must be larger than 0.45 when $\lambda = 0.9999$. If λ is decreased, the required Γ_{init} decreases as well. Figure 9(e) shows that, starting with $\Gamma_{\text{init}} = 0.5$, Γ_{final} equals 0.55, 0.65 and 0.55 for the experiments in which the well-logs are subdivided into $N = 20$, $N = 100$ and $N = 1000$ parts, respectively. Hence, discretisations are biased towards lithology 0, but the discretisation with $N = 100$ is the most strongly biased. This can be explained by a careful study of the vertical transition probabilities: for a fine well-discretisation

Table 2 Coding of lithologies in aquifer system

Sandy lithologies		Clayey lithologies		Other	
Lithology	Code	Lithology	Code	Lithology	Code
Sand and gravel	1	Silt	4	Bedrock	7
Sand	2	Sandy/silty clay	5		
Clayey/silty sand	3	Clay	6		

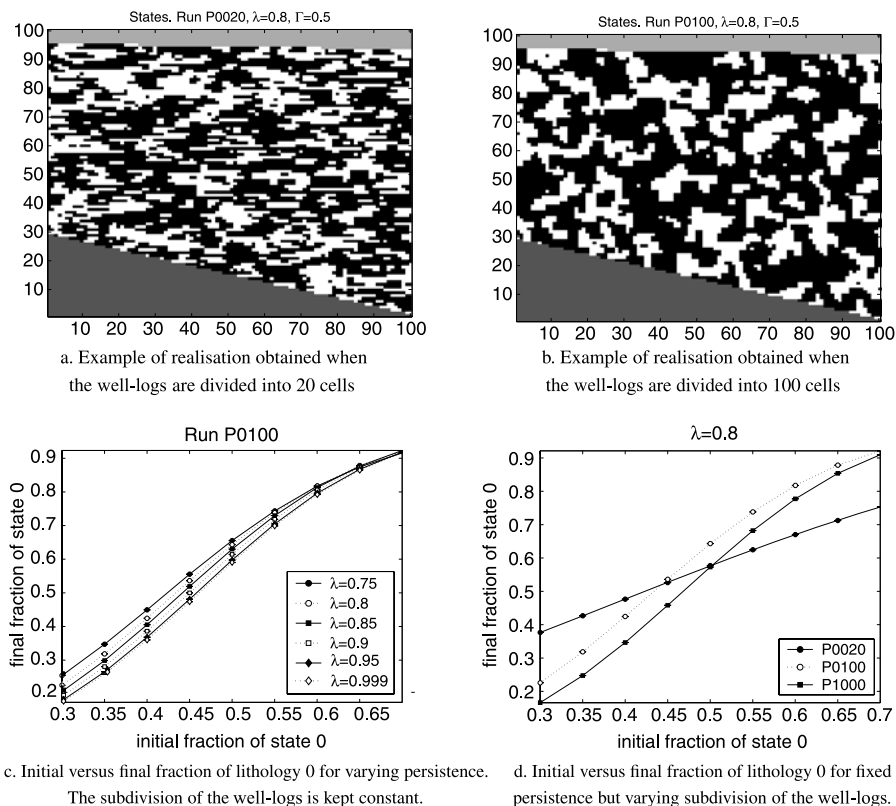


Fig. 9 Results obtained by using well-data

($N = 1000$) there is a slight preference for a transition towards lithology 0, this preference becomes much larger when a coarser well-discretisation is chosen ($N = 100$) and becomes smaller again for a very crude well-discretisation ($N = 20$). A realisation for one-sided probability matrices obtained by subdividing the well-logs in $N = 20$ is shown in Fig. 9(a). Since the single-sided transition probabilities become less diagonally-dominant when the number of well segments is decreased, the cycle length of the obtained solutions and the number of iterations necessary to reach the cyclic solution both increase (not shown).

Seven lithologies In the next experiment, all seven lithologies are considered. Furthermore, the one-sided transition probabilities from air to the various lithologies are taken into account as well. The air–ground interface is fixed. Using this approach, only the lithologies that are at the surface of the wells will be present at the ground–air interface of the complete domain. In the experiment described here, lithologies 3 and 5 are the surface lithologies in the well-logs. In the realisations, only these two lithologies are observed at the surface. Bedrock is taken as an active lithology as well. In the well-logs, either bedrock or lithology 6 is above bedrock and bedrock is always below bedrock. Starting the experiments with prescribed air and bedrock configura-

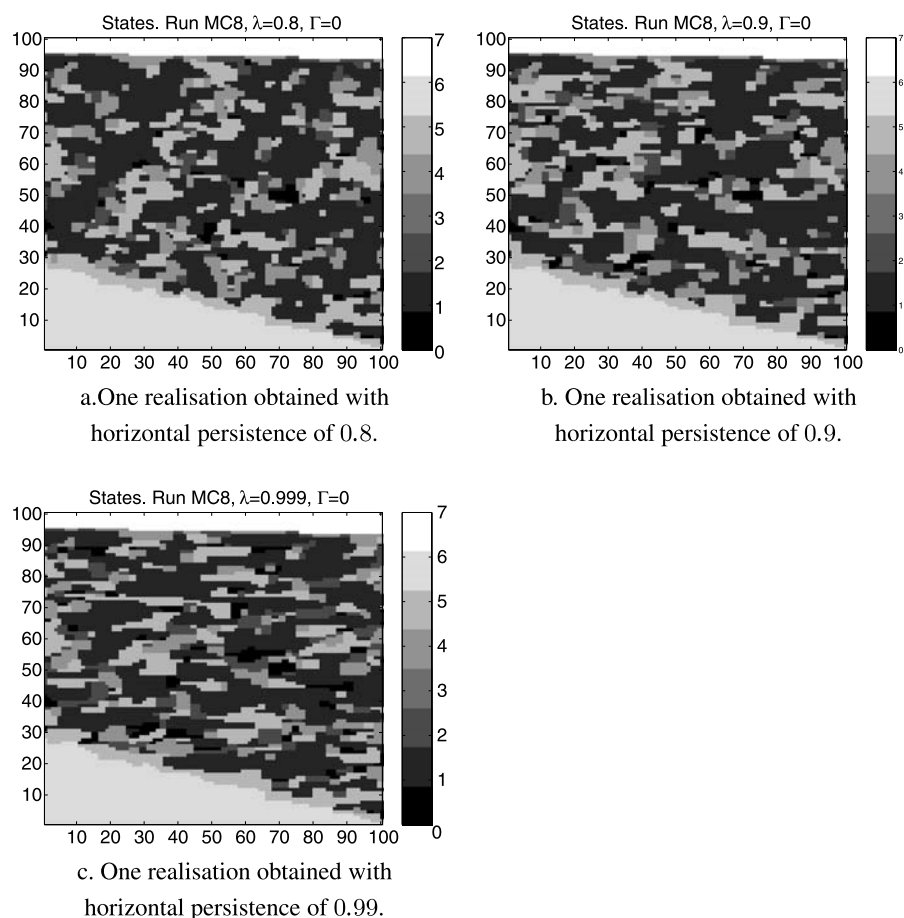


Fig. 10 Realisations obtained with 7 lithologies

tion and initially random choice of lithologies in the intermediate body, the CA-rules result in a configuration with either lithology 6 or bedrock above bedrock and always bedrock below bedrock. The exact interface between bedrock and the other facies is allowed to vary. The results shown in Fig. 10 are obtained with varying horizontal persistence, resulting in a more layered structure when the horizontal persistence is largest. When comparing the model results obtained using the CA-method with those obtained with the coupled Markov chain (CMC) method (Elfeki and Dekking 2001), one can observe the following that with the CMC method only a few lithologies of those present in the well logs are in the final realisations. These are better reproduced with the CA-method. In addition, the CMC method obtains a more layered structure when the cell-size is increased, the same observation can be made for the CA-method when the horizontal persistence is increased. As a final note we would like to emphasise that at this point the rules used in the CA-model are not optimised to reproduce geological structures with a large horizontal extent. The optimisation of CA-rules for this purpose will be the subject of future research.

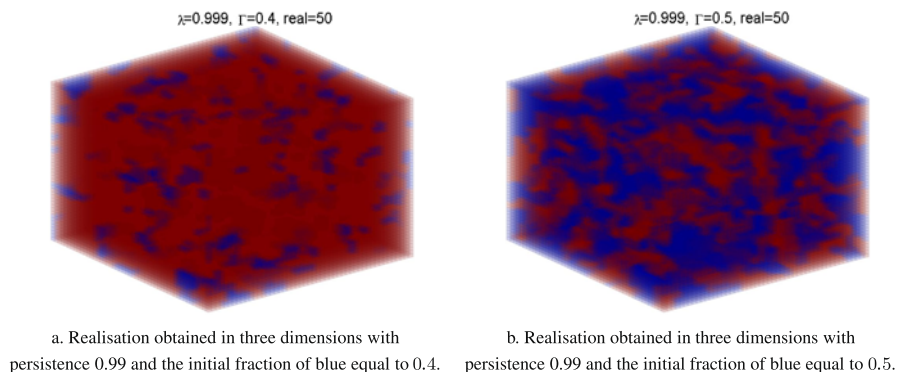


Fig. 11 Realisations obtained in three dimensions

6 Discussion and Conclusions

To characterise the heterogeneity of geological formations, a random CA model is developed. To apply this method, the subsurface is subdivided into N cells, with an initial lithology assigned to each cell. The rules to update the current cell states are chosen from a set of rules, independently for each cell. Following Elfeki and Dekking (2001), the set of rules is obtained by using the one-sided transition probabilities obtained from the well-data, but other schemes are possible (i.e. geological knowledge, training images). In this paper, we focus on the behaviour of the CA model in two dimensions. The number of lithologies is varied between 2 and 8. It is observed that the method converges fast (typically within 50 iterations) to a steady state or periodic solution with a typical cycle length of 10. The realisations within the periodic solutions do not vary much (typically 2–3 percent of the cells change lithology in a periodic solution). Realisations can be easily conditioned on well data. The final fraction of the lithologies can be influenced by choosing different initial fractions. In this way, realisations can be obtained that satisfy geological knowledge about the fraction of the various lithologies in the subsurface. Extension to three dimensions is straightforward. In three dimensions, the model can be conditioned on any number of wells. In Fig. 11, first results are shown for a three-dimensional example: Figure 11(a) for $\lambda = 0.4$ and Fig. 11(b) for $\lambda = 0.5$. Here the subsurface was subdivided into $50 \times 50 \times 50$ cells, 2 lithologies were considered and 2 well-logs were used, positioned in two opposite corners.

The correlation lengths obtained with the model are relatively short. However, right now only the nearest neighbours influence the state of a cell. By increasing the number of neighbours, the correlation length can be extended. Furthermore, one can use training images instead of one-sided transition probabilities to obtain the CA-rules. The influence of the size of the neighbourhood stencil and the choice of the CA-rules is currently under investigation.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- de Marsily G, Delay FF, Goncalves J, Renard P, Teles V, Violette S (2005) Dealing with spatial heterogeneity. *Hydrogeol J* 13:161–183
- Elfeki AAM, Dekking FM (2001) A Markov chain model for subsurface characterization: theory and applications. *Math Geol* 33:569–589
- Elfeki AAM, Dekking FM (2005) Modelling subsurface heterogeneity by coupled Markov chains: Directional dependency, Walther's law and entropy. *Geotech Geol Eng* 23:721–756
- Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22(1):437–467
- Kauffman SA (1984) Emergent properties in random complex automata. *Physica D* 10(1–2):145–156. Cellular automata (Los Alamos, NM, 1983)
- Navoy AS (1991) Aquifer-estuary Interaction and Vulnerability of Groundwater Supplies to sea level rise-driven saltwater intrusion. PhD Thesis, Pennsylvania State University, 225 p
- Strebelle S (2002) Conditional simulation of complex geological structures using multi-point statistics. *Math Geol* 34:1–21
- Toom A (1995) Cellular automata with errors: problems for students of probability. In: Topics in contemporary probability and its applications. *Probab. Stochastics Ser.* CRC, Boca Raton, pp 117–157
- Ulam S (1952) Random processes and transformations. In: Proceedings of the international congress of mathematicians, Cambridge MA, 1950, vol 2. AMS, Providence, pp 264–275
- Wolfram S (2002) A new kind of science. Wolfram Media, Champaign